18

TIME SERIES ANALYSIS

Roland D. Kankey
Assistant Professor
Air Force Insitute of Technology

Somewhere between the simplicity of univariate statistics and the complexities of regression lies the group of estimating techniques used in time series analysis. They range in difficulty from very simple, to fairly complex. All are based upon forecasting using a trend or pattern over time. This article addresses some of the concerns and questions about the use of time series analysis and covers some of the simpler, but useful techniques.

To begin, time series analysis techniques are not deterministic techniques. Unlike regression, they do not reflect a causing relationship. Instead, they all take advantage of a pattern over time. In time series analysis, it is recognized that there are many causing variables. If we expect their magnitudes and effects to remain, we can accept that the net effect of all those variables can be summarized by the changing pattern of cost over time. We thus use time as a stand-in for all the causing variables. We believe that future costs can be predicted based on the pattern formed by past costs.

When this sort of pattern over time has developed, and we believe that all the causing variables will continue to exert the same forces and move along their same paths, then time series analysis techniques provide an excellent alternative for the analyst.

Before using any technique, the analyst needs to have a feel for the situation being estimated. In time series, a critical question is, "Will the trend continue, or will the trend change?" If the analyst feels it will continue, a stable technique should be used. If the trend is likely to change, a responsive technique should be used. The trade off is that a responsive technique reacts more strongly to the expected random variation in costs, resulting in forecasts that "jump around" even if the trend does not change.

Another decision has to be made on what portion of the cost history is appropriate. Graphing the data over time may indicate that the trend has changed in the past. After determining why the trend changed and the likelihood of such changes in the future, the appropriate most recent portion of the data is selected.

Now, what kinds of patterns are possible? The possible patterns (without considering changes in the pattern) include no-trend, linear trend, non linear trend, and cyclic trend. A no-trend pattern is clearly possible over time. In this case, time series may be of limited value in estimating the future. A graph with this type of pattern looks like Figure 1.

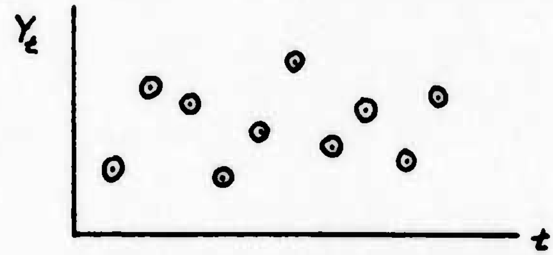


Figure 1: A No Trend Situation

Here we have random variation but an underlying constant value. Time series techniques to forecast under this situation include use of the mean, single moving averages, moving weighted averages, and single exponential smoothing.

A linear trend pattern indicates an underlying linear pattern over time and is illustrated in Figure 2.

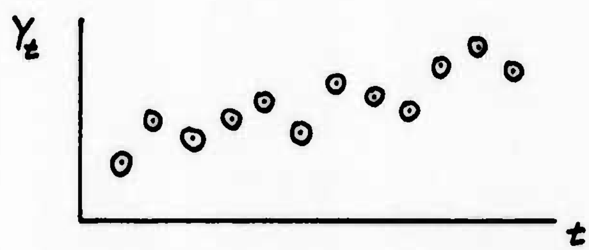


Figure 2: A Linear Trend Situation

If we had experienced such a linear trend in the past and expected it to continue, we would use a linear trend technique such as a least squares best fit line, double moving averages, or double exponential smoothing.

If we had been experiencing a non-linear pattern over time, illustrated in Figure 3, we again have several choices.

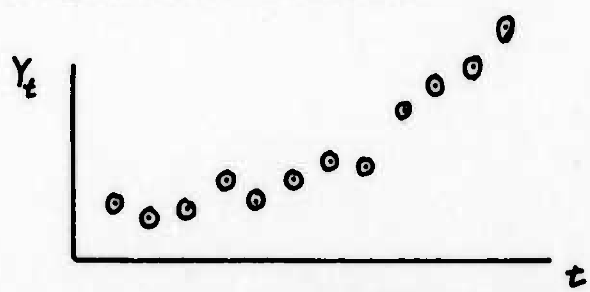
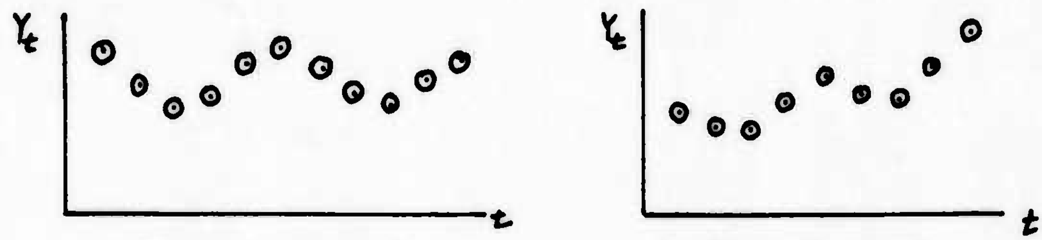


Figure 3: Non-Linear Trend

We can use a linear trend technique and attempt to adjust for the bias (i.e., add or subtract the expected error), or we can use transformations and best fit a non-linear curve to the data. Other choices include more complicated time series approaches, such as triple exponential smoothing, which are beyond the scope of this paper.

Another possible pattern over time is the cyclic pattern, shown in Figure 4:



Cyclic with no trend   Cyclic with trend

Figure 4

In some cases costs will vary over time. For example, seasonal adjustments are made to many government indices and fuel consumption rates (gal/hr) in aircraft vary over the year with higher consumption in hot months and lower consumption in cold. To address this cyclic problem, we can change all the data to cycle-length averages and always forecast an average. This effectively smooths the variations over the cycle. Other, more complicated techniques, are also beyond the scope of this paper.

Now for a brief discussion of some of the basic time series analysis techniques. These basic techniques are those designed for use in no-trend and in linear trend cases. The decision as to whether a no-trend, or a linear trend technique should be used is based upon which type of pattern exists. A look at the data over time may indicate a clear pattern, or it may not. If not, Spearman's Rank test will provide an answer as to which type technique would be more appropriate.

NO TREND

If a no-trend technique is desired and a stable situation exists, then a technique is to forecast future costs as the mean of past costs. If this is felt to be too stable, or not responsive enough, then a single moving average can be used. With single moving averages, the forecast is always the average of the last n data points. The n can be made larger or smaller, depending upon whether more stability or more responsiveness is desired. Using this technique, only the last n data points contribute to the forecast. Each time a new period rolls around and a new data point is developed, an old point is dropped. An embellishment on this technique is the idea of weighted moving averages. Here the analyst has some additional knowledge that the last n data points should not be evenly weighted. For some reason, certain data points should carry more weight than others. A typical weighting procedure would give more weight to the most current data. One procedure, when weighting is felt to be appropriate but no apriori weights are known, is to use the sum of digits approach. This approach for assigning weights says that if we are using n data points we first find 1+2+...+n = k. We then assign weights. The most recent receives a weight of n/k, the next most recent receives a weight of $\frac{n-1}{k}$,... and the oldest a weight of $\frac{1}{k}$. Single moving averages and moving weighted averages are both very easy techniques to understand and use.

A last no-trend technique is called single exponential smoothing. Here, as in each of the previous techniques, a new forecast is made each time period. This technique is similar to the weighted moving average approach but uses a very systematic weighting scheme. In a sense, the single exponential average builds from an assumed starting value and weights each new point with the following procedure:

$$E1\ (c,t) = cY_t + (1-c)\ E1\ (c,t-1)$$

where E1 indicates a single exponential average and E1 (c,t) indicates the single exponential average with a smoothing constant c made at time t.

Y_t is the cost for time period t

E1 (c,t-1) is the previous single exponential average.

It is thus clear that we must start with an assumed exponential average for time t=o and we never throw any data away. It stays in the calculator but is given less and less weight as t gets larger. Two ways to start the process are to let E1 (c,o) = Y_1 or $\bar{Y}$. The preferred method is to use $\bar{Y}$ for the data that is available when you decide to start. Once started, the process is iterative. You do not recalculate E1(c,o) as more data accumulates.

Since c, the smoothing constant, is vital to the process, it's appropriate to look at its characteristics. Its value is bounded by zero and one. Normally the c value is greater than zero and less than .5. The larger the c, the more responsive the technique, the smaller the c, the more stable. It is thus necessary for the analyst to decide whether the pattern is expected to stay the same or change so that an appropriate c can be chosen.

No-trend techniques thus vary from very stable techniques to techniques which can be adjusted for more or less stability/responsiveness. In general, these techniques are designed to be iterated easily as we move through time. All no-trend techniques forecast the future at our current average cost.

LINEAR TREND

If we have a linear trend pattern, than our alternatives include the LSBF line using all the data, double moving averages and double exponential smoothing. The LSBF line uses all the (appropriate) data and is relatively stable. Its forecasts will not be greatly affected by either random variation in the new data point, or a change in trend. It is thus, once again, necessary for the analyst to decide whether the trend is expected to continue or change before choosing a technique. If LSBF is felt to be too stable, i.e., not responsive enough, then another technique should be used.

Double moving averages is a linear trend technique that allows adjustment for more or less stability. The technique uses the following formulas:

$$M1(n,t) = \frac{Y_t + Y_{t-1} + \ldots + Y_{t-n-1}}{n}$$

$$M2(n,t) = \frac{M1(n,t) + M1(n,t-1) + \ldots + M1(n,t-n+1)}{n}$$

M1 (n,t) Single Moving Average at Time t

Y_t The Cost at Time t

M2(n,t) Double Moving Average at Time t

The single and double moving average is calculated for each time period t when data is available. The n is consistent for both and is the number of data points used in each average. Larger n's result in more stable forecasting, smaller n's make the technique more responsive. When using this linear trend technique the forecast is in the form of a straight line ray with this ray used to forecast future costs. The forecast equation made at time t for the future involves a_t and a b_t where:

$$a_t = [2\ M1(n,t)] - M2(n,t)$$

$$b_t = \frac{2}{n-1}\ [M1(n,t) - M2\ (n,t)]$$

The forecasting equation is then

$$FM2(n,t,t+i) = a_t + b_t * i$$

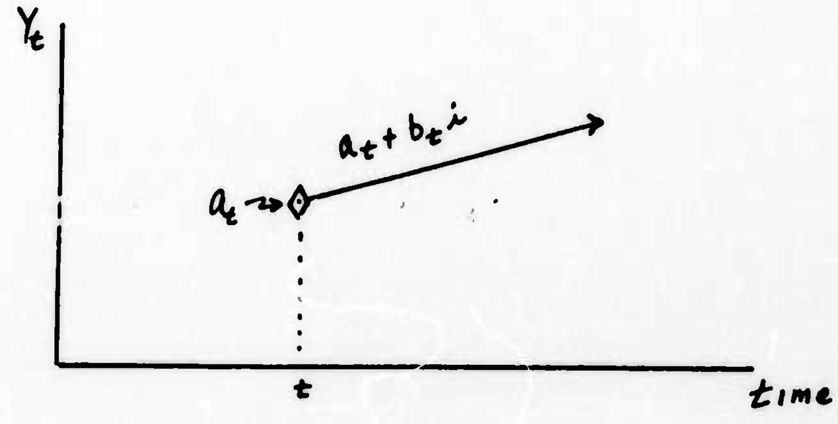
FM2 Our forecast using double moving averages

n The number of items averaged

t The time period when the forecast is made

i The numbers of periods ahead for the forecast (the forecast interval)

If we graph this ray we have:



The data points used for a double moving average forecast are the last 2n-1 points.

Another linear trend technique is double exponential smoothing. This technique is an expansion of the single exponential smoothing and likewise uses a very systematic weighting process. The double exponential average is:

$$E2(c,t) = cE1(c,t) + (1-c)\ E2(c,t-1)$$

While the single exponential average remains:

$$E1(c,t) = cY_t + (1-c)\ E1(c,t-1)$$

and:

E2 indicates a double exponential average

E1 indicates a single exponential average

t indicates the time period when the averages are made

c indicates the smoothing constant

Y_t indicates the cost at time t

This process is very easy to iterate. For example, if we find that the cost for time 101 was 44, we can iterate from the data:

t	Y_t	E1(.2,t)	E2(.2,t)
100	40	38	36
101	44		

To get the new single and double exponential averages:

$$E1(.2,101) = (.2)(44) + (.8)(38) = 39.2$$

$$E2(.2,101) = (.2)(39.2) + (.8)(36) = 36.64$$

Since this is a linear trend technique, we will once again be forecasting costs on a ray defined by a_t and b_t where they are now defined as:

$$a_t = 2E1\ (c,t) - E2\ (c,t)$$

$$b_t = \frac{c}{1-c}\ (E1(c,t) - E2(c,t))$$

Our forecast equation for i periods ahead is:

$$FE2\ (c,t,t-i) = a_t + b_t * i$$

If we use the example's data to forecast 4 periods ahead:

$$a_{101} = 2(39.2) - 36.64 = 41.76$$

$$b_{101} = \frac{.2}{.8}(39.2 - 36.64) = .64$$

$$FE2\ (.2,101,105) = 41.76 + .64 * 4 = 44.32$$

The technique of using exponential averages to forecast costs is called exponential smoothing.

One point on double exponential smoothing that should be discussed is "How do we start?" As shown, each exponential average is a weighted average of all prior data. How did we calculate an E1 (c,0) and E2(c,0) so that the iterative process can begin? The recommended procedure is to do a LSBF on the data available when you want to start the process. Set your regression equation $a = a_t$ and $b = b_t$, then work the formulas in reverse to solve for an E1(c,0) and E2(c,0). Note that since we are now in a linear trend situation , the continued use of $\bar{Y}$ as our first E1(c,0) would be illogical. Some will, however, indicate that how you start is not critical since the process that assigns weights will assign less and less weight to that assumed value. Exponential smoothing is a child of the computer age. It is easy to iterate but it assumes a reasonable amount of data, say 15 data points (more or less data needed depending upon the smoothing constant chosen).

SELECTING A TECHNIQUE

Now, how do we evaluate the forecasting ability of a particular technique in a particular case. We do not have any probability assumptions and cannot compare F tests or Prediction Intervals. What we can do is evaluate how well the technique as performed in the past. To do this we make forecasts for previous time periods and compare those forecasts with the actual reported costs. Since we do not expect such forecasts to exactly equal the costs we need to evaluate the errors.

The most useful measure of these errors is the Mean Absolute Deviation (MAD)

$$MAD = \frac{\sum_{terms} |ERRORS|}{\text{Number of Terms Considered}}$$ or the average of the absolute errors.

If the MAD is 2.13, we can say that on average the technique's forecasts are off by 2.13 units. In general, a technique with a smaller MAD, based on past (simulated) forecasting would be seen as a better choice to forecast the future; if the technique is acceptable in terms of stability and responsiveness.

No technique is always useful, that is why the analyst needs to know several different techniques. Time series techniques are simpler to use than regression techniques and help fill the gap between statistics and regression. If a time series technique is appropriate and gives satisfactory results, it should be used.